AN INTERMEDIATE SEMANTIC REPRESENTATION EXTRACTED FROM ENGLISH TEXT FOR SIGN LANGUAGE GENERATION

Eva Safar, Ian Marshall

School of Information Systems
University of East Anglia
es@sys.uea.ac.uk, im@sys.uea.ac.uk

Abstract. Natural sign languages have a number of similarities to oral natural languages, though the three dimensional nature of the space around a signer offers a number of opportunities unavailable to oral languages. In this paper we describe some aspects of a possible route for the translation between the two modalities from English text to British Sign Language (BSL). Using an example we focus on the extraction of an intermediate semantic representation as Discourse Representation Structure (DRS) from English text and how this semantic structure is related to the syntax of a sign language in the HPSG framework.

1. Introduction

Only in the last half century have sign languages been recognized as 'real' languages leading to the increased recognition of the status and the needs of the deaf people in society. Thus sign languages are nowadays accepted as minority languages, which coexist with the majority language [Neidle et al. 2000] and more importantly that deaf people prefer signed languages as their first language are the motivation for providing increasing support and services for the Deaf in their own native language instead of a written form of English. Therefore the ViSiCAST project focuses on the use of virtual humans (avatars) [Kennaway 2001], which present deaf signing via a semi-automatic translation from English to a sign language gesture notation.[1]

2. Sign Language

We give a short description of the major features of sign languages based on [Brien 1992] and [Sutton-Spence and Woll 1999] using the following example.

(1) `I take some mugs from the table.`

This sentence is realized with a different structure to English word order. BSL has a topic-comment structure, in which the main informational subject or topic is signed first, in this case 'table' and 'mug'. However 'i' is signed last as it has a pronominal realization,

[1]This work is incorporated within *ViSiCAST*, an EU Framework V supported project which builds on work supported by the UK Independent Television Commission and Post Office. The project develops virtual signing technology in order to provide information access and services to Deaf people.

which equals pointing in sign language. The verb is an agreement verb, which incorporates information about person and number of the subject and indirect object. This is realized by the direction of the movement of the verb in syntactic signing space. In this case the sign for the verb begins at the location of the object and ends at the location of the subject. Because of this agreement the subject can be dropped (prodrop). The plural can be expressed by the repetition of the movement (usually having a distributive meaning, in contrast to a sweep movement which means collective plural). Some verbs allow a handshape to be incorporated within the sign of the verb, in this case the form of the 'mug' as we hold it. It also has been noted how crucial the positioning of objects is, as in this case the mugs are on the table (before they are moved), therefore the 'table' and the 'mugs' have to be signed in a close physical position. Facial expressions associated with the position of eyebrows distinguish between declarative (neutral brows), yes/no questions (raised brows) and wh-questions (furrowed brows). Mouth patterns can help disambiguate manually similar signs ('table' vs 'flat'). Facial expression and body posture can provide adverbial information (like intensity) and indicate the signer's attitude to the accompanying proposition. In sign language research a so called glossing is used to transcribe signing (XA and Y standing for locations and (3) signalling 3 repetitions):

(2) `TABLE-XA MUG TAKE-XA-Y(3) I-Y`

3. Semantic Representation Formalism

In Section 2. example (2) gives a sentence gloss, which shows similarities to logical representations of that sentence. Discourse Representation Structures (DRS) [Kamp and Reyle 1993] have been chosen because of further considerations of sign language phenomena. DRSs allow isolation of tense/aspect and modifying phenomena that are realized in different sign language grammatical constructs or modalities. This means that rather than express temporal information by morphological or syntactic features associated with verbs, it is expressed with the help of time lines in the signing space, by the ordering of the propositions, temporal adverbials or even facial expressions. In addition, the centrality of co-referentiality in DRSs reflects the need to appropriately determine how to assign fixed positions in signing space to significant discourse referents[2].

In [Kamp and Reyle 1993] event propositions are labeled for use as arguments with temporal predicates. The labeling is a very helpful method which allows a flat representation of the hierarchical structure of arguments and operator embeddings. Since sign languages also reflect distinctions of manner and degree by using facial expressions modifying the meaning of the main proposition and head nod for negation, the labeling has been extended by introducing labels for each semantic predicate. Labeling has also been used in Vermobil [Dorna and Emele 1996; Dorna 2000], but in contrast to its uniform labeling an ontology for all DRS propositons has been introduced to ease the mapping from the flat semantic structure of the DRS to the input structure of the target language specific HPSG.

[2] We note however, that DRSs have to be supported by other pronoun resolution algorithms, to choose the appropriate referent when more than one is available. The intersentential pronoun resolution is therefore supported by a modified version of [Kennedy and Boguraev 1996] algorithm, which includes a database of female and male names.

4. Extraction of Semantics from Syntax

In the first step the English sentence is processed via the CMU parser [Sleator and Temperley 1991] that generates the following linkage structure for our sentence:

```
(3)      +---------------------Xp---------------------+
         |           +-------MVp-------+              |
         |           +-----Op----+     +----Js---+    |
         +--Wd-+-Sp*i+     +-Dmc-+     |   +--Ds-+    |
         |     |     |     |     |     |   |     |    |
       ///// I.ps take.v some mugs.n from the table.n .
```

From different options given by the CMU parser we chose the approprate one containing an 'MVp' link (the prepositional phrase as verb modifier) rather than 'Mp' (which would signify a modification to the noun). 'Wd', 'Op' and 'Js' indicate nouns with different syntactic functions (subject, object and noun in prepositional phrase respectively), 'D' stands for determiner, where 'p' and 's' indicate number. These links are put in linear order according to their relational representation as a collection of n-tuples:

```
(4)
 [[m, 1, 0, Wd, Wd, Wd, 1], [m, 1, 1, Sp*i, Sp*i, Sp,2],[m,1,3,D,Dmc, Dmc, 4],
  [m, 2, 2, O, Op,Op,4],[m, 3, 2, MV, MVp, MVp, 5], [m, 1, 6, D, Ds, Ds, 7],
  [m, 2, 5, J, Js, Js, 7]]
```

This list is then processed via Definite Clause Grammar (DCG) rules implemented in Prolog. A link dictionary (see a simplified example for an entry below) maps each link type to a λ-expression DRS definition (λ-DRS) [Marshall and Safar 2001]. When the stem of a word has to be found, WordNet [Miller et al. 1993] is consulted and the extracted morphological (temporal and/or number) information is used in the DRS[3].

```
(5)    dict(det,Sent, Links,[_,_,W1,_,Dlink,_,W2],E) :-
         (matchUC(Dlink,'D'),E=lambda(P,lambda(Q,merge(merge(drs([X],[]),P@X),Q@X))).
```

The DCG then concatenates the λ-DRSs in the appropriate order and instantiates the arguments of the predicates appropriately [Blackburn and Bos 1999]. Functional application (β-reduction) and DRS merge operation combine the λ-DRSs into the final DRS [Bos et al. 1994]. The final DRS looks like the following:

```
(6)       sentence1
            drs(0):[v(0), v(1)],
              [a(0):i(v(0)),
                  drs(1):[v(2)],
                    [q(0):some(v(2)),
                     a(1):mug(v(2)),
                     c(0):plural(v(2))]
                             >      drs(2):[],
                                       [t(0):when(e(0)),
                                        t(0)=now,
                                        e(0):take(v(0), v(2)
              a(2):table(v(1)),
              l(0):from(e(0), v(1))]
```

[3]We plan to use WordNet more widely in our future work for word sense disambiguation and definite descriptions resolution.

This representation is consistent with [Kamp and Reyle 1993] but has been extended, as mentioned in Section 3., by labeling each semantic predicate (a = nominal, e = event, t = temporal, l = local, q = quantification, c = number, v = variable) and also instead of a diamond notation for the quantifier 'some', we introduce it in the condition part of the DRS followed by '>' as symbol for *implies*.

5. From Semantics towards the Syntax of Sign Language

The input SEM structure is converted from the above DRS representation (example (6)) to a nested HPSG semantic structure (example (7)) as required by the ALE generation algorithm [Carpenter and Penn 1999].

```
(7)       (sent,sem:(index:Ind1,mode:decl,
                  restr:[(sit:Ind3,reln:from,figure:a, ground:b,
                          args:[(index:a,
                                 restr:[(sit:Ind3,reln:take,act:c,thm:d,
                                         args:[(index:c,
                                                restr:[(sit:Ind3,reln:i)]),
                                               (index:d,
                                                count:(number:pl,collordist:distrib),
                                                restr:[(sit:Ind3,reln:mug)])])]),
                                (index:b, count:number:sg,
                                 restr:[(sit:Ind3,reln:table)])])]))
```

The conversion uses the ontology of DRS propositions introduced by the labeling, In our example the mapping from the the DRS embedding of the event predicate and the referent of the location in the location predicate (label l) is clearly mirrored in the SEM structure. The verb also reflects the two argument predicate of the DRS's e(0) predicate, while the number information is extracted via the c(0) label supported by the DRS embedding for distributive meaning.

This structure contains enough information to start the generation process using a sign language HPSG [Safar and Marshall 2002]. Much of the detail of its feature structure is focused on fine grain detail in the PHON component describing how signs are constructed from handshape, palm orientation, finger direction and movement information. This information uses a notation called HamNoSys [Prillwitz et al. 1989]. The argument structure, locations and the agreement components of the SYN structure determine conditions under which signs can be combined into a grammatically correct physical realisation (as described in Section 2.) using the information from the input SEM structure embeddings, indexing and count feature derived from English syntax via the DRS.

Example (8) shows a typical SYN feature structure using the lexical entry 'take'. The left hand side (LHS) of the lexical entry consists of three parts: 1. the gloss, 2. the non-manual components and 3. the manual components including the mouthing at the front in SAMPA notation. The uninstantiated variables stand for the orientation of the hand (Efd) and the start and end positions (Heightobj, Distobj, Heightsubj, Distsubj) of the movement, which depend on the location of the object that is moved from one position to the other. In our example the mug is moved from the table to the speaker. R1 represents a possible pluralisation. The value of the 'allow_pl_repeat' feature allows pluralisation, so if the object is plural, R1 is instantiated to a HamNoSys symbol that represents a repeated movement to signal distributive meaning of the plural. Similarly 'allow_pl_sweep' would allow a sweeping movement to signal collective meaning if its value was not negative. The 'gref' feature determines whether the subject or the object agrees with the verb in number. Hsh (handshape) and Plm (palmorientation) are associated with the object noun (precomp(lement)s), which is informed of a manipulation based handshape. The order of precomps is vital for the surface realization of arguments. In 'context' the 'add_list' feature determines the positions of the nouns in signing space. Variable 'Gloss' stands for the object 'mug', that is allocated two locations, as it is moved from one position to another. One of those positions has to be the same as the subject's position, because this is the goal or final position. 'Cin' is the allocation map that contains the free and occupied positions of the signing space. The final map (Cout) cannot contain two locations for the 'mug' as the source positions is no longer valid. Therefore the 'delete_list' determines which position has to be freed. Sign languages allow pronoun drop, whenever the handshape, the starting and the end position of the verbal sign includes one or two of the complements. This phenomenon reflects a similar relation between rich agreement and non-overt expression of subject/object pronomina as in many languages, such as Hungarian or Italian. The 'pro_drop' features specify whether nominals are obligatorily or optionally omitted ('must' or 'can' values).

```
(8)   [[take],[Brow],['teIk',Nhd, Hsh, Efd, Plm, Heightobj, Distobj, R1, hamreplace,
                     Efd, Plm, Heightsubj, Distsubj]]  --->
        syn:(precomps:[(@ nmanip(Ph,Gloss,Index2,Precomp1,Hsh,Efd,Plm,Nhd,Sg)),
                     (@ np2(W,Glosssubj,Plm2, EfdT,Index1,Precomp2,Num,PLdistr))],
            postcomps:[],
            allow_pl_repeat:yes_loc_indiv_finite,
            allow_pl_sweep:no,
            head:(dirmanipverb_lxm,
                 agr:(num:(number:sg,collordist:Coll),
                      per:per,gref:Index2),
                 context:(context_in : Cin,
                          add_list   : [(glossref:[(ref:Index1,glossr:Glosssubj)],
                                              locat:locatefd:Efdsubj,
                                              distance:Distsubj,
                                              heights:Heightsubj),
                                         (glossref:[(ref:Index2,glossr:Gloss)],
                                              locat:locatefd:Efdsubj,
                                              distance:Distsubj,
                                              heights:Heightsubj),
                                         (glossref:[(ref:Index2,glossr:Gloss)],
                                              locat:locatefd:Efd,
                                              distance:Distobj,
                                              heights:Heightobj)],
                          context_out: Cout,
                          delete_list:[(glossref:[(ref:Index2,glossr:Gloss)],
                                              locat:locatefd:Efd,
                                              distance:Distobj,
                                              heights:Heightobj)]),
                 prodrp_subj:(first:can,
                         second:cant,
                         third:cant),
                 prodrp_obj:(first:can,
                         second:cant,
                         third:cant))),
        phon:phon, sem:sem.
```

6. Conclusions

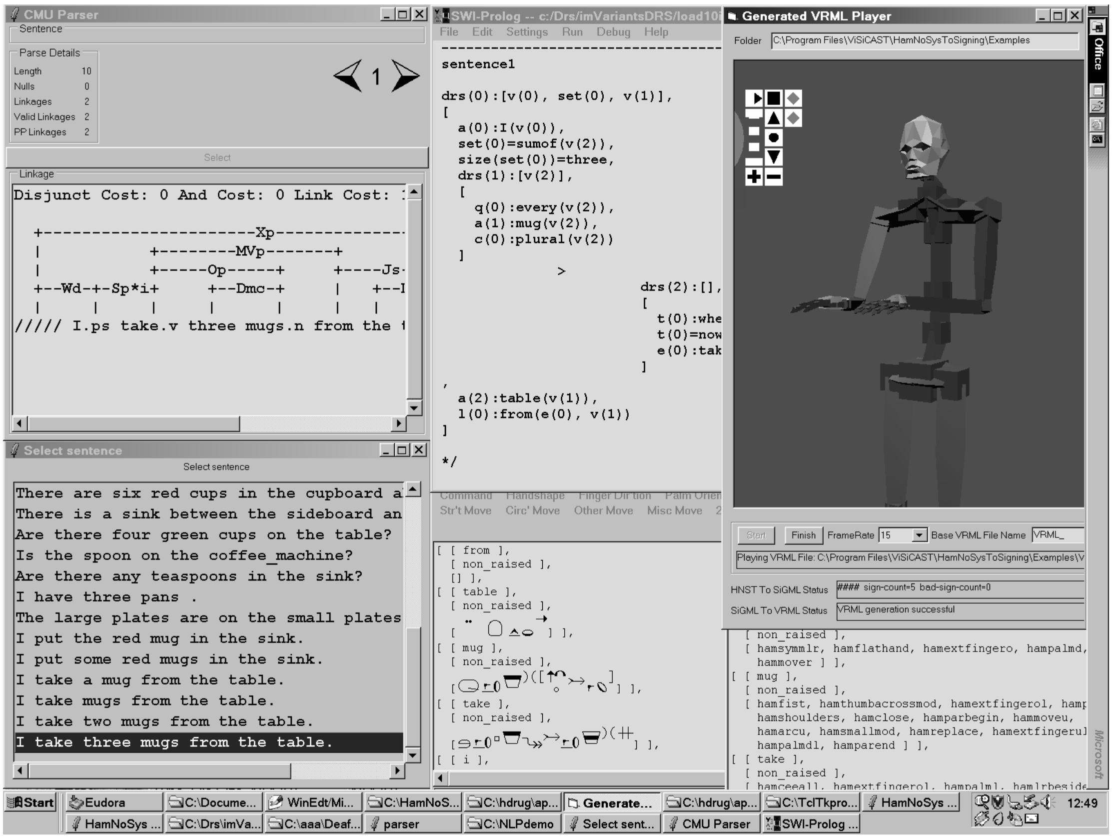

Figure 1: Screenshot of the current translation system

In this paper we have shown a possible route for the translation from English text to British Sign Language (BSL) focusing on the extraction of an intermediate semantic representation as Discourse Representation Structure (DRS) from English text that is suitable for successful and relatively straightforward conversion into a nested semantic input for a sign language HPSG syntactic generator. Figure 1. illustrates the current demonstrator system that allows selection of a sentence which is passed to the CMU parser and then to the DRS generator and HPSG synthesis systems.

References

Blackburn, P. and Bos, J.: 1999, *Representation and Inference for Natural Language. A First Course in Computational Semantics. Volume II.*, http://www.coli.uni-sb.de/~bos/comsem/book1.html

Bos, J., Mastenbroek, E., McGlashan, S., Millies, S., and Pinkal, M.: 1994, *A Compositional DRS-based Formalism for NLP Applications. Report 59*, Universitaet des Saarlandes, Saarbruecken

Brien, D. E.: 1992, *Dictionary of British Sign Language/English*, Faber and Faber, London,Boston

Carpenter, B. and Penn, G.: 1999, *The Attribute Logic Engine. User's Guide. Version 3.2 Beta*, Bell Labs

Dorna, M.: 2000, Semantikbasierter transfer - maschinelle uebersetzung auf mengen linguistischer constraints. (doctoral thesis), in *Arbeitspapiere des Instituts fr maschinelle Sprachverarbeitung (AIMS). Vol.6, No.2*

Dorna, M. and Emele, M.: 1996, Semantic-based transfer. report 143., in *Proceedings of the 16th International Conference on Computational Linguistics (COLING-96)*, Copenhagen

Kamp, H. and Reyle, U.: 1993, *From Discourse to Logic. Introduction to Model theoretic Semantics of Natural Language, Formal Logic and Discourse Representation Theory*, Kluwer Academic Publishers, Dordrecht

Kennaway, J.: 2001, Synthetic animation of deaf signing gestures, in *The Fourth International Workshop on Gesture and Sign Language Interaction, Gesture Workshop 2001 (GW2001)*, City University, London, UK

Kennedy, J. and Boguraev, E.: 1996, Anaphora for everyone: Pronominal anaphora resolution without a parser, in *Proceedings of the 16th International Conference on Computational Linguistics COLING'96*, Copenhagen, Denmark

Marshall, I. and Safar, E.: 2001, Extraction of semantic representations from syntactic cmu link grammar linkages, in *Proceedings of Recent Advances in Natural Lanugage Processing*, Tzigov Chark, Bulgaria

Miller, G., Beckwith, R., Fellbaum, C., Gross, D., and Miller, K.: 1993, *An Introduction to WordNet. An On-line Lexical Database*, http://www.cogsci.princton.edu/~wn/

Neidle, C., Kegl, J., MacLaughlin, D., Bahan, B., and Lee, R.: 2000, *The Syntax of American Sign Language*, MIT Press

Prillwitz, S., Leven, R., Zienert, H., Hanke, T., Henning, J., et al.: 1989, *Hamburg Notation System for Sign Languages - An Introductory Guide*, International Studies on Sign Language and the Communication of the Deaf, Volume 5., Institute of German Sign Language and Communication of the Deaf, University of Hamburg

Safar, E. and Marshall, I.: 2002, Sign language translation using drt and hpsg, in *A. Gelbukh (Ed.) Procieedings of the Third International Conference on Intelligent Text Processing and Computational Linguistics , CICLing, Mexico, Lecture Notes in Computer Science 2276*, Springer Verlag, Mexico

Sleator, D. and Temperley, D.: 1991, *Parsing English with a Link Grammar*, Carnegie Mellon University Computer Science technical report CMU-CS-91-196

Sutton-Spence, R. and Woll, B.: 1999, *The Linguistics of British Sign Language. An Introduction*, University Press, Cambridge